



STICHTING KIJKONDERZOEK

---

# DUTCH VAM

## MANDATORY CALCULATION AND REPORTING RULES

MANDATORY RULES AND RECOMMENDED PROCEDURES



**STICHTING KIJKONDERZOEK**  
**MANDATORY CALCULATION AND**  
**REPORTING RULES**  
MANDATORY RULES AND RECOMMENDED  
PROCEDURES

AMSTERDAM 1-6-2016

**THIS DOCUMENT IS WRITTEN BY SKO, IN COLLABORATION WITH TNS NIPO, KANTAR MEDIA  
AUDIENCES AND NIELSEN**

© AUTHOR'S COPYRIGHT.

Any unauthorized use of the document subsequent to its publication in an electronic or other medium Any distribution or publication of said document to others who obtain copies of it from by any means without written authorization by SKO shall constitute an infringement of the Author's copyright.



## INDEX

<b>INTRODUCTION</b>	<b>4</b>
<b>1 MANDATORY CALCULATION RULES</b>	<b>8</b>
<b>1.1 AUDIENCE RATING</b>	<b>9</b>
1.1.1 ONLINE BROADCAST DAY / TOTAL / 28	9
1.1.2 VIEWING MOMENT REPORTS	10
1.1.3 ALL PUBLISHERS	11
1.1.4 ONLINE COMMERCIALS RATING	12
1.1.5 ONLINE COMMERCIALS VAST_TOTALS	13
<b>1.2 REACH</b>	<b>13</b>
1.2.1 PROGRAM REACH (BROADCAST DAY/TOTAL/27)	13
1.2.2 PUBLISHER REACH (EITHER BROADCAST OR VIEWING MOMENT REPORTS)	14
1.2.3 CAMPAIGN REACH	15
<b>1.3 VIEWING TIME</b>	<b>16</b>
<b>1.4 RESULTS FOR MULTIPLE VIDEOS</b>	<b>17</b>
1.4.1 CUMULATIVE AUDIENCE RATING	17
1.4.2 CONTACT FREQUENCY DISTRIBUTION	17
1.4.3 CUMULATIVE CONTACT FREQUENCY DISTRIBUTION	18
1.4.4 CUMULATIVE REACH	18
1.4.5 AVERAGE CONTACT FREQUENCY	19
<b>1.5 SAMPLES</b>	<b>19</b>
1.5.1 DAILY SAMPLES	19
1.5.2 PERIOD SAMPLES	20
1.5.3 WEIGHT FACTOR	21
1.5.4 TARGET GROUP SELECTION	22
1.5.4.1 TARGET GROUP SELECTION FOR DAY – OR AVERAGE RESULTS	22
1.5.4.2 TARGET GROUP SELECTION FOR PERIOD SAMPLES (APPLICABLE TO REACH)	22
1.5.5 ROUNDING OFF	24
1.5.6 POPULATION SIZE	24
1.5.7 CONVERTING PERCENTAGES TO ABSOLUTE NUMBERS	24
<b>2 ADDITIONAL MANDATORY CALCULATION RULES</b>	<b>25</b>
2.1 SUMMER- AND WINTER TIME	26
<b>3 MANDATORY REPORTING RULES</b>	<b>28</b>
<b>3.1 SOURCE</b>	<b>28</b>
<b>3.2 REPORT TYPES</b>	<b>28</b>
<b>3.3 PERIOD IN THE YEAR</b>	<b>28</b>



<b>3.4</b>	<b>TIME SLOT</b>	<b>29</b>
<b>3.5</b>	<b>TARGET GROUP</b>	<b>29</b>
<b>3.6</b>	<b>SAMPLE SIZE</b>	<b>29</b>
<b>3.7</b>	<b>REPORTING CRITERIA SKO-VIM</b>	<b>29</b>
3.7.1	REPORTING CRITERIA MINIMUM REACH	29
3.7.2	REPORTING CRITERIA GRPS	30
<b>3.8</b>	<b>POPULATION SIZE</b>	<b>30</b>
<b>3.9</b>	<b>ROUNDING OFF RESULTS</b>	<b>30</b>
<b>4</b>	<b>RECOMMENDED RULES</b>	<b>32</b>
<b>4.1</b>	<b>RECOMMENDED RULES SKO TOTAL VIEWING</b>	<b>32</b>
4.1.1	RESULT TYPES USED BY SKO	32
4.1.2.	OTHER SKO RECOMMENDATIONS	33
<b>APPENDIX 1</b>	<b>DEFINITIONS</b>	<b>34</b>
<b>1.</b>	<b>AUDIENCE RESEARCH FILES</b>	<b>35</b>
<b>1.1</b>	<b>RESPONDENT DATA FILES</b>	<b>35</b>
1.1.1	VIEWING STATEMENTS	35
1.1.2	BACKGROUND CHARACTERISTICS	35
1.1.3	WEIGHT FACTORS	36
<b>1.2.</b>	<b>BROADCAST DATA FILES</b>	<b>37</b>
1.2.1	PROGRAMS	37
1.2.2	CLASSIFICATION OF BROADCAST COMMERCIALS	38
1.2. 3.	ONLINE COMMERCIALS	38
1.2.4.	ONLINE ONLY VIDEO CONTENT	39
<b>2.</b>	<b>CALCULATION OF THE LENGTH OF THE VIEWING STATEMENT OF ONLINE PROGRAMMES FILES</b>	<b>39</b>
<b>APPENDIX 2</b>	<b>PERIOD SAMPLE CRITERIA</b>	<b>40</b>
<b>APPENDIX 2</b>	<b>OVERVIEW OF VARIABLES NOT FOR EXTERNAL REPORTING</b>	<b>44</b>



## INTRODUCTION

SKO (Stichting KijkOnderzoek), the Dutch television audience research foundation, has now a secondary dataset available on the online viewing behaviour in the Dutch population (OAM). This is next to the existing TAM dataset that is provided to the market.

SKO now provides clients with Online Audiences Data “raw data” on a daily basis, starting October 1<sup>st</sup> 2015 for online video-content. For online commercials data is provided per June 1<sup>st</sup> 2016. Raw data refers here to the validated viewing records on the respondent level for video-content and online commercials, as well as files containing weight factors and respondent background information, device information and (broadcast and commercial) metadata information. SKO provides a file format description and the guidelines for software used for processing the raw data.

The new dataset can be used to calculate 5 types of viewing reports:

1. Online Linear Viewing (**OLV**); the online linear viewing during broadcast (live viewing)
2. Online Broadcast Day (**OBD**); the online viewing on the day of broadcast (linear viewing + VOSDAL)
3. Online Broadcast Total (**OBT**); the online viewing on the day of broadcast and the 6 consecutive days (linear viewing + UGK6)
4. Online Broadcast Total 28 days (**OB28**); the online viewing on the day of broadcast and the 27 consecutive days (live + UGK27)
5. Online Viewing Moment (**OVM**); viewing within a specific time frame or period in time (e.g. day/week/month) of available online video content, unrelated to broadcast moment, date of publication nor origin.

In this report, three types of rules are described:

- Mandatory calculation rules
- Mandatory reporting rules
- Recommended rules

### **Mandatory calculation rules:**

- Each user of the survey data is required to follow the SKO standard calculation rules for all 4 viewing types, in order to prevent different results being used in the market. The SKO guidelines are intended to insure that the calculations of the various result types are carried out in the same manner, regardless of the party carrying them out.
- SKO provides basically two sets of calculation rules: for the report types ‘Online Broadcast Day/Online Broadcast Total(28 days)’ and ‘Viewing moment’ which users of the survey data are required to use.



### **Mandatory reporting rules:**

- In addition to the calculation rules, SKO has established reporting rules that users of the raw or aggregated data are required to follow when reporting or publishing results.

### **Recommended rules:**

- In addition to the mandatory calculation and reporting rules, SKO has developed recommended procedures to determine less central result types. These are intended to further increase agreement in the market.

### **The basic report types:**

Many of the calculations have great similarities with the way TAM results are calculated. Where possible we refer to those calculations that are already widely available in existing systems and software. The main differences are in the selection of data to calculate reports on.

### **Data within the data set:**

The imported/processed dataset should hold 5 'data buckets'. Data buckets classify the data in terms of content viewed and link to broadcast day metadata.

1. **Linear viewing**; viewing a linear online stream of a broadcaster (either broadcaster or cable operator owned player)
2. **VOSDAL, viewing on the same day as live, is TSV on the same day as broadcast, but not linear viewing**
3. **TSV 1-6 days**; viewing content 1 day after StartDAY (DAY 0) and Broadcast DAY  $\leq$  6 days.
4. **TSV 7- 27 days**; viewing content after 6 days Broadcast DAY StartDAY (DAY 0)  $>6$  and BroadcastDAY  $\leq$  27 days.
5. **Other**; viewing content that is online only (not related to an specific broadcast), or TSV  $>$  27 days of broadcast day.

### **Building report types from data buckets**

1. Online Linear Viewing (**OLV**); Bucket 1 **Linear viewing**
2. Online Broadcast Day (**OBD**); Bucket 1 **Linear viewing** AND Bucket 2 **VOSDAL**
3. Online Broadcast Total (**OBT**); Bucket 1 **Linear viewing** AND Bucket 2 **VOSDAL** AND Bucket 3 **TSV 1-6 DAYS**
4. Online Broadcast Total 28 days (**OB28**); Bucket 1 **Linear viewing** AND Bucket 2 **VOSDAL** AND Bucket 3 **TSV 1-6 DAYS** AND Bucket 4 **TSV 7-27 DAYS**
5. Online Viewing Moment (**OVM**); Bucket 5 **Other  $\geq$  27 DAYS** AND content in **bucket 1 to 4 at the moment of viewing**.

The last reporting option asks for a data structure where, for each viewing session, it is possible to know when the viewing was done (viewing moment), and to what broadcast moment the viewing relates.



For online commercials data, viewing moment is always applicable as the viewing is always at the moment of play out.

TABLE 1.1 REPORT TYPES AND SKO RULES

Report types	Linear viewing/ Online Broadcast Day/Total/28	Viewing moment reports
<b>Behaviour measured</b>	Viewing at the moment of broadcast (day 0) and Viewing at some time other than the moment of broadcast on the day of broadcast (day 0) or on one of the six, or 27 following days	Viewing that takes place within a given time frame, regardless the content origin. Or if it concerns online commercials.
<b>Calculation</b>	See Chapter 1	See Chapter 1
<b>Reporting</b>	See Chapter 2	See Chapter 2

*The dark grey cells in the table indicate that mandatory rules apply. SKO will advise in a later version on using guidelines for result and report types other than those mandatory.*

In Table 1.2, the mandatory rules and recommended procedures for the calculation of various result types for different report types are given. As in Table 1.1, dark grey indicates that mandatory rules apply; light grey indicates result types in the various reports for which SKO provides recommended procedures.

TABLE 1.2 REPORT TYPES AND RESULT TYPES

Report type:	Linear viewing/ Online Broadcast Day/Total/27	Viewing moment reports
<b>Calculation</b>		
Audience rating	√	√
Market share	X	X
Reach	√	√
Viewing time	√	√
Results for multiple pieces of content	√	√
Samples	√	√
<b>Reporting</b>		
User reports of all result types	√	√
Reporting of all result types for publishing	X	X
<b>Recommendations</b>		
Profile	√	√
Top list	√	√



Overlap/ Duplication	√	√
Selectivity index	√	√
Population sizes	√	√

There are no separately defined guidelines for various result types in the set of reporting rules. The reporting rules must be followed in the reporting for all results.

*The dark grey cells in the table indicate that mandatory rules apply. SKO also advises using the guidelines laid out in this document for result and report types indicated in the light grey cells.*

*√ = SKO has established a rule for this result type in the report type concerned*

*X = SKO advises against using this result type in the report type concerned*

*To avoid ambiguity in the audience figures, SKO requires that what is being reported and what the basis of outcomes is be precisely indicated.*





CHAPTER 1

# **MANDATORY CALCULATION RULES**



## 1 MANDATORY CALCULATION RULES

### 1.1 AUDIENCE RATING

#### 1.1.1 ONLINE BROADCAST DAY / TOTAL / 28

The audience rating (kdh) for a program is defined as the average percentage of viewers per second during the program. Respondents in the sample are weighted with a factor determined by the extent to which their background characteristics are represented in the total population in the Netherlands.

$$kdh_i = \frac{\sum_{r=1}^R w_r p_{r,i}}{\sum_{r=1}^R w_r} \cdot 100\%$$

$kdh_i$  : audience rating for online program  $i$

$R$  : number of respondents in the sample

$w_r$  : weight for  $r$

$p_{r,i}$  : viewing probability for respondent  $r$  for program  $i$

Where viewing probability is given by

$$p_{r,i} = \frac{K_{r,i}}{dur_i}$$

$K_{r,i}$  : duration of viewing of respondent  $r$  for online program  $i$

$dur_i$  : duration of online program  $i$

Please take into account that duration of the online programme refers to the clip duration (sko\_cl) or the net duration of the TV broadcast in case of programmes (excluding breaks) provided in the Nielsen files as well as the net duration of the TV-breaks.

Two types of programs are distinguished:

1 = Linear programme or Programme previously broadcasted (< 27 or >27 days after initial broadcast day)

2 = Online only content

The type 1 programmes can be used to either calculate Online Broadcast Day (OBD) / Online Broadcast Total (OBT) / Online Broadcast Total28 days (OB28). The calculation method is the same



for all three result types. The difference is in the viewing data (bucket/ type of online content) that is taken into account.

OBD: all viewing statements are to be taken into calculation, that happen on the initial broadcast day. Either linear or time shifted day 0.

OBT: all viewing statements are to be taken into calculation, that happen on the initial broadcast day plus the consecutive 6 days.

OB28: all viewing statements are to be taken into calculation, that happen on the initial broadcast day plus the consecutive 27 days.

Multiple sorts of programs types can be found in the Nielsen PRL files (see Appendix 1 definitions). Linear streaming related to TV broadcast contains two types of programs:

- Programmes
- Breaks

Advertising breaks should always be considered as online programmes when calculating ratings.

Averages over multiple online programmes should be calculated as follows: a file is calculated on a daily basis for the specific target group containing the audience rating per online programme, rounded off to six decimal places precisely. This data is then used to calculate the averages over broadcasts

#### 1.1.2 VIEWING MOMENT REPORTS

The average audience in the VMR analysis (Viewing Moment Reports) is calculated in an identical way as for TV broadcast content in the calculations described before. The only difference is that for all content viewed calculations are to be done on those viewing statements within the reporting framework. So we always look at the viewing statements 'at the moment of consumption'. Irrespective if it relates to a broadcast on another day. So it basically is reporting within a specific time slot, based on the viewing statements that happened within that time slot. So that can be derived from all data buckets (1-5).

Viewing moment is the only report applicable to online commercial data, as the viewing is always at the moment of play out.

The audience rating for a time slot (time interval) is defined as the average percentage of viewers per second during the time slot. The starting and ending times for the interval are not rounded down.



$$kdh_t = \frac{\sum_{r=1}^R w_r p_{r,t}}{\sum_{r=1}^R w_r} \cdot 100\%$$

$kdh_t$  : audience rating for time interval  $t$

$R$  : number of respondents in the sample

$w_r$  : weight for respondent  $r$

$p_{r,t}$  : viewing probability of respondent  $r$  for time interval  $t$

Averages per time slot over several days should be calculated as follows: a file is calculated on a daily basis for the target group concerned containing the audience rating per publisher/content selections per minute, rounded off to six decimal places precisely. This data is then used to calculate the averages over several days, rounded off to the desired number of decimal places.

$$\overline{kdh}_t = \frac{\sum_{d=1}^D kdh_{t,d}}{D}$$

$\overline{kdh}_t$  : average audience rating for time interval  $t$  over several days

$D$  : number of days over which the average is calculated

$kdh_{t,d}$  : audience rating for time interval  $t$  on day  $d$

All periods are calculated on the basis of all individual days. In a quarterly averaging, for example, the month of January counts somewhat more heavily than the month of February.

### 1.1.3 ALL PUBLISHERS

Total Online Audience rating is defined as the average percentage of viewers per second watching online video during a time slot. This should be always calculated as a viewing moment report. See 1.1.2

$$kdh_{z,t} = \sum_{z=1}^Z kdh_{z,t}$$

$kdh_{z,t}$  : audience rating all publishers for time interval  $t$

$kdh_{z,t}$  : audience rating for publishers  $z$  in time interval  $t$

$Z$  : total number of publishers



#### 1.1.4 ONLINE COMMERCIALS RATING

The audience rating of a TV commercial is defined as the average percentage of viewers per second to the station broadcasting the spot during the starting minute of the spot. As for online commercials the data contains only the % seen of the commercial, it is needed to calculate viewing time estimates for online commercial viewing distributions.

An online campaign can have one of the following results on respondent level (only one value per respondent will be present in the data set):

- 0-viewing
- Start
- 25% complete
- 50% complete
- 75% complete
- 100% complete.

The actual length of the online commercial is provided in the harmonized Nielsen Online Commercial files (.COL).

It is needed to use the %seen to calculate a synthetic viewing statement. This should be done in the following manner, stepwise:

From the Nielsen metadata (COL) get the actual length of the commercial. E.g. 20 seconds.

0-viewing -> nil viewing

- If a respondent has no start statement, this needs to deliver null viewing.
- Start - > If a respondent has a start statement this should be calculated as 5% of the length of the commercial, with a minimum of 1 second.
- If a respondent has a 25% seen statement, this needs to deliver a 37.5% length viewing statement.
- If a respondent has a 50% seen statement, this needs to deliver a 62.5% length viewing statement.
- If a respondent has a 75% seen statement, this needs to deliver a 87.5% length viewing statement.
- For 100% it delivers a 20 second statement.

All seconds need to be rounded down if not an equal number (e.g. 5.5 secs becomes 5).

The result is a viewing estimate for the length of the online commercial.

Once these statements have been calculated, the calculation below can be used.



$$kdh_s = \frac{\sum_{r=1}^R w_r p_{r,s}}{\sum_{r=1}^R w_r} \cdot 100\%$$

$kdh_s$  : audience rating for spot  $s$

$R$  : number of respondents in the sample

$w_r$  : weight for respondent  $r$

$p_{r,s}$  : viewing probability of respondent  $r$  for the length of the online commercial  $s$

### 1.1.5 ONLINE COMMERCIALS VAST\_TOTALS

It is not allowed to calculate output that refers to 'Total starts' or the like. The basic information on the use of online commercials comes from the VAST census measurement. The metrics from that research should also be calculated as described in the calculation and reporting rules for VAST Online Commercials Census data. For each element that is reported based on the panel data (i.e. per publisher or brand), starts can be also calculated on the VAST OC Census data. To be reported alongside the panel output.

## 1.2 REACH

### 1.2.1 PROGRAM REACH (BROADCAST DAY/TOTAL/27)

Program reach is defined as the percentage of viewers that had watched one program from a series of programs within a specific time interval at least one time for a specified amount of time. A single-online broadcast or program is considered a special case of a program series: a series of one.

No lower limit for program reach has been set to be included in reach. So any viewing is enough to be accounted for as being 'reached'. One is free to specify conditions for reach such as the number of online programs watched in the series or time interval, the amount of time that must have watched, and whether viewing had to be uninterrupted. The length of time someone must have watched one or more programs can be expressed in minutes or as a percentage of program length.

All of the conditions used in determining the reported program reach must be specified in the reporting.



$$reach_I = \frac{\sum_{r=1}^R w_r d_{r,I}}{\sum_{r=1}^R w_r} \cdot 100\%$$

$$d_{r,I} = \begin{cases} 1 & j_{r,I} \geq J \\ 0 & j_{r,I} < J \end{cases}$$

- I* : series of contacts with online programs within a specific time frame or interval  
*d<sub>r,I</sub>* : 'dummy' indicating whether respondent is included in reach according to specified definition  
*J* : criterium for the specified program reach  
*R* : number of respondents in the sample  
*w<sub>r</sub>* : weight factor for respondent *r*  
*j<sub>r,I</sub>* : personal criterium of respondent *r* for *I*, calculated according to specified conditions

#### 1.2.2 PUBLISHER REACH (EITHER BROADCAST OR VIEWING MOMENT REPORTS)

Publisher reach is defined as the percentage of viewers that watched a specific publishers' content for a specified length of time at least once during a series of time slot. A single time slot is treated as a special case of a series of time slots.

No lower limit for program reach has been set to be included in reach. So any viewing is enough to be accounted for as being 'reached'. The user is free to specify whether a respondent must have watched during one or more time slots in order to be included in the specific publisher reach, how long he or she must have been watching and whether or not this was uninterrupted.

All of the conditions used in determining a reported publishers' reach reported must be specified in the reporting.



$$reach_{z,T} = \frac{\sum_{r=1}^R w_r d_{r,z,T}}{\sum_{r=1}^R w_r} \cdot 100\%$$

$$d_{r,z,T} = \begin{cases} 1 & K_{r,z,T} \geq X \\ 0 & K_{r,z,T} < X \end{cases}$$

- $T$  : series of contacts with online content within a specific time frame or interval  
 $reach_{z,T}$  : *publishercontent* reach for publisher  $z$  during the series of time intervals  $T$   
 $R$  : number of respondents in the sample  
 $w_r$  : weight of respondent  $r$   
 $K_{r,z,t}$  : personal criterium of respondent  $r$ , calculated according to the specified conditions for publisher  $z$  in  $T$   
 $X$  : criterium for the specified publisher reach

### 1.2.3 CAMPAIGN REACH

The calculation of spot reach is a yes/no calculation of reach and frequency. In the data there are The following options/levels of reach thresholds. The user should be able to choose any of the thresholds. SKO advices to use 25% as the standard reach threshold, bit this is not mandatory:

- Start
- 25% complete
- 50% complete
- 75% complete
- 100% complete.

After defining the threshold reach can be calculated as it is being done for time interval or frame (see 1.3.2). A split per day/week and or publisher/device should be possible.





$$reach_I = \frac{\sum_{r=1}^R w_r d_{r,I}}{\sum_{r=1}^R w_r} \cdot 100\%$$

$$d_{r,I} = \begin{cases} 1 & j_{r,I} \geq J \\ 0 & j_{r,I} < J \end{cases}$$

- $I$ : series of contacts with online campaigns within a specific time frame or interval  
 $d_{r,I}$ : 'dummy' indicating whether respondent is included in reach according to specified definition  
 $J$ : criterium for the specified campaign reach  
 $R$ : number of respondents in the sample  
 $w_r$ : weight factor for respondent  $r$   
 $j_{r,I}$ : personal criterium of respondent  $r$  for  $I$ , calculated according to specified conditions

Warning: in the data only the highest value is available so if you want to compute reach it should be taken into account that if a respondent has a value 'complete' he/she also counts as reached for the other (lower) criteria.

### 1.3 VIEWING TIME

Viewing time is defined as the average time that viewers watched during a program or, within a time slot, a publisher, expressed in minutes.

$$kt_t = \frac{\sum_{r=1}^R w_r K_{r,t}}{\sum_{r=1}^R w_r}$$

- $kt_t$ : viewing time for time interval  $t$   
 $R$ : number of respondents in the sample  
 $w_r$ : weight for respondent  $r$   
 $K_{r,t}$ : viewing time for respondent  $r$  for time interval  $t$

Reporting on viewing time per time slot must be based on the viewing records available in the raw data.



Please take into account that duration of the online programme refers to the clip duration (sko\_cl or the net duration of the TV broadcast in case of programmes (excluding breaks) provided in the Nielsen files as well as the net duration of the TV-breaks.

#### 1.4 RESULTS FOR MULTIPLE VIDEOS

##### 1.4.1 CUMULATIVE AUDIENCE RATING

The cumulative audience rating is defined as the audience rating summed over multiple pieces content being watched.

$$ckdh = \sum_{i=1}^n kdh_i$$

*ckdh*: cumulative audience rating for *n* online programs

*n*: number of online programs

*kdh<sub>i</sub>*: audience rating for online program *i*

The cumulative audience rating for spots generates the number of GRPs, as one GRP is defined for spots as 1% rating (*kdh*).

##### 1.4.2 CONTACT FREQUENCY DISTRIBUTION

The contact frequency distribution is defined as the distribution of contacts based on reach, expressed as a percentage of persons per number of contacts.

$$f(c, K) = \frac{\sum_{r=1}^R w_r f_r(c, K)}{\sum_{r=1}^R w_r} \cdot 100\% \quad \text{voor } c = 0, 1, \dots, K$$

*f(c, K)*: percentage persons that had *c* contacts with online commercials or online programs within the defined time frame or interval *K*

*w<sub>r</sub>*: weight for respondent *r* probability that respondent *r* had *c* contacts on the basis of *K*

Contact frequency distribution can be calculated for programs and spots. It is important here that that the definition of contact used in the above formula is consistent with the calculation model used. For spots and programs the contact is 0 (not reached) or 1 (reached), in accordance with the definition of program- or spot reach (see § 1.3.1 and 1.3.3).



#### 1.4.3 CUMULATIVE CONTACT FREQUENCY DISTRIBUTION

Cumulative contact frequency distribution is defined as the distribution of contacts based on reach, expressed as a percentage of persons per number of contacts. This result is calculated on the basis of a normal frequency distribution by summing the results. The group 1+ is made up of the groups 1-1, 2-2, n-n.

A formula for the percentage of persons that have seen between the  $c$  and  $C$  online commercials or programs within the defined time frame is:

$$f(c, C, K) = \sum_{i=c}^C f(i, K)$$

$f(c, C, K)$ : percentage of persons that have seen between  $c$  and  $C$  online commercials or programs  
 $f(i, K)$ : percentage of persons that  $i$  online commercials or programs have seen

By selecting  $C=K$  in the above formula,  $c+$  can be calculated:

$$f(c+, K) = \sum_{i=c}^K f(i, K)$$

$f(c+, K)$ : percentage of persons that have seen at least  $c$  online commercials or programs within a time frame or interval  
 $f(i, K)$ : percentage of persons that  $i$  online commercials or programs have seen

#### 1.4.4 CUMULATIVE REACH

Cumulative reach is defined as the percentage of persons reached after a series of events. This is the equivalent of cumulative contact frequency distribution for the group 1+. However, it is easier to calculate with the formula below:

$$cb = \frac{\sum_{r=1}^R w_r \cdot (1 - \prod_{k=1}^N (1 - p_{r,k}))}{\sum_{r=1}^R w_r} \cdot 100\%$$

$cb$ : percentage of persons reached after a series of  $N$  online commercials or programs within a time frame or interval  
 $w_r$ : weight for respondent  $r$   
 $p_{r,k}$ : 'contact' probability of respondent  $r$  for online commercial or program  $k$

The formula is universally valid, provided the definition of contact probability used is in agreement with the relevant reach model (see also § 1.3.1 and 1.3.3.).



#### 1.4.5 AVERAGE CONTACT FREQUENCY

Average contact frequency (gcf) is defined as the average number of contacts that reached viewers who had been reached.

$$\bar{c} = \frac{\sum_{c=1}^K c \cdot f(c, K)}{\sum_{c=1}^K f(c, K)}$$

$\bar{c}$  : average number of contacts for respondents who had at least one contact

$f(c, K)$  : percentage of persons that had  $c$  contacts after  $K$  broadcasts

The above formula can be used for online programs as well as online commercials.

$$\bar{c} = \frac{\sum_{c=1}^K c \cdot f(c, K)}{\sum_{c=1}^K f(c, K)} = \frac{\sum_{i=1}^K kdh_i}{\sum_{c=1}^K f(c, K)}$$

$\bar{c}$  : average number of contacts for respondents that had at least one contact

$f(c, K)$  : percentage of persons that had  $c$  contacts after  $K$  online commercials or programs

$kdh_i$  : audience rating for online commercial or program  $i$

If the reporting period span more than one day, then the audience ratings used in the formula must be calculated using a period sample. This is in contrast to the reported audience ratings for the individual days, which are calculated based on the daily sample.

In both of the above formulas, the denominator is the cumulative or 1+ reach.

### 1.5 SAMPLES

#### 1.5.1 DAILY SAMPLES

The daily samples are used for all results that cover a single day, as well as for averages over several days and online programs.

The daily sample contains all of the panel members in those households that were successfully contacted on the day, that means all those whose characteristics are contained in the demographic files in the raw data.



### 1.5.2 PERIOD SAMPLES

Period samples are used for all results that cover one or more days (with the exception of averages over several days or online programs).

The period for period samples always spans one or more whole weeks, which begin on Monday and end on Sunday. The first step in setting up the period sample is to determine in which weeks the earliest and the most recent broadcasts occurred. The period is bounded by the Monday of the earliest week and the Sunday of the most recent week and includes all the days in between.

To include panel members in period samples, two criteria must be met, based on the presence of data for individual panel members in the demographic files in the raw data. The table in Appendix 3 gives: the maximum number of a specific day of the week for which a respondent's data can be missing, and the maximum total number of days of data that can be missing.

In the exceptional case that the most recent week has not yet ended, it will be assumed that contact for each respondent for the missing days of the last week was successfully made.

Because of concerns on reliability, samples spanning more than five quarters cannot be used. These criteria for period samples are used for the calculations of the report types.

**Weeks:** The number of weeks in the period sample.

**Week days:** The maximum number of a specific day of the week that can be missing

**Total:** The maximum number of days in total that can be missing

Weeks	Week-days	Total
1	0	0
2	1	1
3	1	1
4	1	2
5	1	2
6	1	2
7	1	2
8	1	3
9	1	3
10	1	3
11	1	3
12	1	3
13	1	4
14	1	4
15	1	4

Weeks	Week-days	Total
28	2	8
29	2	9
30	2	9
31	2	9
32	2	10
33	2	10
34	2	10
35	2	11
36	3	11
37	3	11
38	3	12
39	3	12
40	3	12
41	3	12
42	3	13



Weeks	Week- days	Total
16	1	5
17	1	5
18	1	5
19	1	6
20	1	6
21	2	6
22	2	6
23	2	6
24	2	7
25	2	7
26	2	8
27	2	8

Weeks	Week- days	Total
43	3	13
44	3	13
45	3	14
46	3	14
47	3	14
48	3	15
49	3	15
50	3	15
51	4	15
52	4	16
53	4	16
54	4	16

### 1.5.3 WEIGHT FACTOR

The weight factor for each respondent in the period sample is calculated by averaging the daily weightings for that respondent. The total number of days in the period sample – and not merely the number of days on which the broadcasts in question took place - is used as the divisor for all respondents in the sample.

#### **BE AWARE:**

The online panel data is calibrated to census and consist of expanded panel data. It basically means that every record will be in the file for an average of about 10 times. By that a virtually 50.000 panel is delivered. Each weighting factor will be 1/10<sup>th</sup> of the initial one. Securing that the total reporting panel still delivers a total population rating. You are to use *the 000.0000-999.9999 weight factor from the data* included in the viewing records (AUvX150809\_calibrated.DEM).

The average weight factor is multiplied by a correction factor to compensate for the loss of respondents in the period sample, so that the sum of the weightings remains equal to the total population (projection). The correction factor is equal to the sum of the daily weightings of all respondents, including those ultimately not selected, for the day in question and for all of the days in the period sample, divided by the weightings of all of the selected respondents in the period sample.



$$w_r = \left( \frac{\sum_{d=1}^D \sum_{r=1}^{R_d} \frac{w_{r,d}}{D}}{\sum_{d=1}^D \sum_{r=1}^{N_d} \frac{w_{r,d}}{D}} \right) \cdot \frac{\sum_{d=1}^D w_{r,d}}{D} = \left( \frac{\sum_{d=1}^D \sum_{r=1}^{R_d} w_{r,d}}{\sum_{d=1}^D \sum_{r=1}^{N_d} w_{r,d}} \right) \cdot \frac{\sum_{d=1}^D w_{r,d}}{D}$$

- $w_r$  : Period weight factor for respondent  $r$
- $w_{r,d}$  : Weight factor for respondent  $r$  on day  $d$
- $D$  : Number of days in period sample
- $R$  : Total number of respondents in the daily sample
- $N$  : Number of respondents included in the period sample

In the case that the most recent week in the period is not yet ended, the averaging of weightings is carried out to and including the last completed day (and thus not the Sunday of the most recent week).

#### 1.5.4 TARGET GROUP SELECTION

Target groups can be defined based on the demographic data supplied in the raw data.

The values for the demographic characteristics from the base survey are updated twice a year (on Monday week 1 and on Monday of week 27) for all respondents. New demographics can be also introduced on this dates, if applicable.

A number of demographic are included in the calibration process. A short list of targets used in calibration is provided in the documentation on data integration on the SKO website.

##### 1.5.4.1 TARGET GROUP SELECTION FOR DAY – OR AVERAGE RESULTS

For target selection in a day sample it important to use the demographic file of the specific day when the results are calculated. In case that averages over several days are calculated, demographic files of these specific days should be used.

##### 1.5.4.2 TARGET GROUP SELECTION FOR PERIOD SAMPLES (APPLICABLE TO REACH)

When period samples are used, the demographic data used in the target group selection should be taken from the day in which the broadcasting of/over 80% of the period, campaign or broadcasts is reached.

That means that the demographic characteristics should be use for the day in which the 80% of broadcasting, schedule or day in the period is reached.

For example, if a schedule includes 7 broadcasts in a period of to , then the period (the date) of/over 80% of the broadcasts should be used. In this examples, this is day 11.



Example:

7 broadcasts in a period of 2 weeks

day	Broadcast	%
1	1	14%
2		
3	2	29%
4		
5	3	43%
6		
7	4	57%
8		
9	5	71%
10		
11	6	86%
12		
13	7	100%
14		



To determine the sample period please refer to the calculation rule # 1.6.2. If a respondent falls within the period sample but it was not present in the demographic file of the specific day, the demographic characteristics of this particular respondent should be looked for in the demographic files of the previous day (day -1, day -2, ...).

For example, if a schedule includes 7 broadcasts in a period of to , then the period (the date) of/over 80% of the broadcasts should be used. If the respondent was not present in the demographic file of this specific day (day 11), then the demographic file of the previous day (day 10) should be used for this particular respondent.

Example:

7 broadcasts in a period of 2 weeks

day	Broadcast	%
1	1	14%
2		
3	2	29%
4		
5	3	43%
6		
7	4	57%
8		
9	5	71%
10		
11	6	86%
12		
13	7	100%
14		



Respondent not present in demo file







#### 1.5.5 ROUNDING OFF

All calculations are carried out without rounding off.

If interim result types are calculated, they must be rounded off to six decimal places, to insure a negligible level of discrepancy with regards the reference. In the final reporting, results are rounded off to one decimal place. Weightings are always rounded off to one decimal place.

#### 1.5.6 POPULATION SIZE

The population size for the 53 commercial target groups is reported annually by SKO on the Monday of week 1. These are based on the most recent figures the MOA Golden Standard. The updated results from MSS will be updated on the Monday of week 5.

#### 1.5.7 CONVERTING PERCENTAGES TO ABSOLUTE NUMBERS

The population sizes for the 53 commercial target groups established annually by SKO should be used when converting percentages to absolute numbers (e.g. audience rating, reach).

When determining absolute numbers for other target groups, established population sizes from other sources should be used.

If population size is to be estimated based on the audience research, using the daily sample for the first day that changes in population sizes or demographics become effective is recommended.

- When an analysis straddles the changeover dates (Monday of week 1 for the 53 SKO commercial target groups, Monday of week 1, 27 for non-standard target groups), the population numbers applicable on the day should be used per day for results based on daily samples.
- For results based on period samples and for analysis that straddles the changeover dates, men should use the population sizes that applied to the period in which the broadcasting of/over 80% of the broadcasts is reached.

In principle, the calculation of absolute numbers should be carried out based on percentages that have not been rounded off. If necessary, however, percentages may be rounded off to six decimal places exactly.



CHAPTER 2

# **ADDITIONAL MANDATORY CALCULATION RULES**



## 2 ADDITIONAL MANDATORY CALCULATION RULES

### 2.1 SUMMER- AND WINTER TIME

On the first day of the summer time period (Sunday, 02:00 hours becomes 03:00 hours) information of the first hour (02:00 - 03:00) is not included neither in the viewing statements (swo\_wwwwd.DAT) nor in The Nielsen Company registrations and reporting because this hour does not exist.

On the Sunday of the changeover from summer time to winter time, one hour (02:00:00 –02:59:59) occurs twice. The viewing statements (SWO\_wwwwd.DAT) of this additional hour are present in the Kantar data, where an hour is deducted from the viewing statements of the additional hour. Therefore, statements with a start time at 01:00 hours and end time till 02:00 hours are present in the data. In order to avoid incorrect ratings, the software neglects the additional hour. For the time registration in the Nielsen files timing proceeds as usual for the preceding Saturday through 25:59:59; on Sunday, timing begins at 02:00:00 hours.



CHAPTER 3

# **MANDATORY REPORTING RULES**



### 3 MANDATORY REPORTING RULES

This chapter contains the rules SKO has set governing the reporting of results of the video audience survey, which all users of the audience survey data are required to follow. As such, they can be considered as an extension of the calculations rules.

The calculation rules discussed above must be followed in any reporting of results from the video audience research. In reporting results from the video audience research the following rules must be followed:

#### 3.1 SOURCE

“Stichting KijkOnderzoek’ or ‘SKO’ must be given as the (data) source in any reporting of the results of the video audience measurement.

The report and result type, time of year, the time slot or program and the target group must also be specified when publishing results.

#### 3.2 REPORT TYPES

SKO requires that the following nomenclature be used in the publishing reportage: Result type (=xxxx).

1. Online Broadcast Day (**OBD**); the online viewing on the day of broadcast (live+VOSDAL)
2. Online Broadcast Total (**OBT**); the online viewing on the day of broadcast and the 6 consecutive days (live+UGK6)
3. Online Broadcast Total 28 days (**OB28**); the online viewing on the day of broadcast and the 27 consecutive days (live+UGK27)
4. Online Viewing Moment (**OVM**) reports; viewing within a specific time frame or period in time (e.g. day/week) of available online video content, unrelated to broadcast moment, date of publication nor origin.
5. Online campaign results.

An overview of variables that can and cannot be reported on, based on the available data will be provided later in June. Be aware that for example the reporting based on the panel data is not allowed on creativeID, only on OnlineID and ItemID as the lowest level.

#### 3.3 PERIOD IN THE YEAR

The period of the year can be indicated by year, quarter, month, week or day. A starting and ending date must always be given for non-standard periods..



### 3.4 TIME SLOT

This refers to the portion of the day from which the data is drawn and is indicated in hours and minutes. For example, '16.00 to 18.00' indicates that the results apply to the interval from 16.00 hours up to and including 17.59 hours. For campaigns the minimum period is a complete day.

### 3.5 TARGET GROUP

This is the group of persons to whom the result type applies. The target group must always be indicated. This should be done using the socio-demographic variables and product-use variables defined in the codebook provided by SKO.

### 3.6 SAMPLE SIZE

The unweighted sample size must be specified.

### 3.7 REPORTING CRITERIA SKO-VIM

#### 3.7.1 REPORTING CRITERIA MINIMUM REACH

SKO has established that ratings should be published only when they are reliable. For the SKO-VIM data, this means that only results on reliable reach may be published. This is a condition for reliable result reporting.

In order to test whether the figures can be published for a specific target group, we compare these with a net reach threshold. Net reach is calculated in terms of the total number of unweighted panel members being reached in the daily or period sample. The minimum reach of 60 (unweighted) individuals for the target group to be reported is demanded, next to a total panel members reached of a minimum of 100.

***So for reporting ratings for a target group for programmes, a minimum of 60 panel members (unweighted) reached must be met, and on the total 6+ group this must be a minimum of 100 (unweighted) panel members.***

If the number of (unweighted) members reached does not reach the 60 persons in target group rule, results are only to be reported for the total audience at 6+ target level.



For online commercials results the reporting rule is that ***a minimum of 60 panel members (unweighted) reached in the reporting target must be met.***

***If the threshold is not met, in display of results they should be flagged or marked by a RED colour.***

The reach threshold of 60/100 or 60 persons in the sample applies to previously specified reach conditions. For instance, if the lower limit for reach has been set to include viewers that watched at least 60 consecutive seconds of a time slot, reach can only be reported reliably if at least 60/100 respondents viewed 60 consecutive seconds of the time slot. If the reach threshold is set as the half of the program length, non-consecutive viewing, thereach may be reported when a minimum of 60/100 respondents has viewed (non-consecutive) at least half of the programme. The reporting software must flag if the numbers are not met.

### 3.7.2 REPORTING CRITERIA GRPS

When the reach of a campaign, or a selection of days within the campaign, meets the publication criteria, the corresponding Online GRPs and related result types may be reported on a (measurable) reliable basis.

When the reach of commercial campaigns fails to reach the threshold, the GRP's results may not be published. Still the VAST events can be reported as totals.

### 3.8 POPULATION SIZE

As stated in § 1.7.4, population sizes for 53 standard target groups are reported annually by SKO. When reporting population sizes for other target groups, the source for the determination of these population numbers must be given.

SKO advises against using population sizes other than those published by SKO for the 53 standard target groups. The sample is not weighed for characteristics other than those that form the standard groups. If the population size is to be estimated based on the audience research, then SKO recommends using the daily sample for the first day that changes in population sizes or demographics become effective for the estimation.

### 3.9 ROUNDING OFF RESULTS

With respect to rounding off, the following apply in all reportage:

- absolute numbers should be published as whole thousands,
- percentages should be published to a maximum of one decimal place.



CHAPTER 4

# RECOMMENDED RULES





## 4 RECOMMENDED RULES

### 4.1 RECOMMENDED RULES SKO TOTAL VIEWING

In addition to the mandatory calculation and reporting rules, SKO has developed recommended procedures for defining less central results. They are intended to increase agreement in the market.

#### 4.1.1 RESULT TYPES USED BY SKO

The result types defined in this section are employed by SKO on the web site and in publications such as the annual report.

##### **Profile**

The profile gives the composition of the audience with respect to one or more relevant variables, expressed in percentages. Percentages are usually given for the target group TOT 6+ or TOT 13+. The term 'viewer' can be defined based on audience ratings or reach. The most clear-cut profile is one with mutually exclusive categories that together total 100%.

For example: 160 people aged 6 years and older watched a broadcast. This group is composed of 120 women and 40 men. The profile for this broadcast for the variable sex is thus: women 75% and men 25%.

##### **Top list**

SKO uses the following rules for top lists (details available at:

[https://kijkonderzoek.nl/images/OTV\\_Methods/150217\\_SKO\\_Calculation\\_and\\_reporting\\_rules\\_OTV\\_raw\\_data.pdf](https://kijkonderzoek.nl/images/OTV_Methods/150217_SKO_Calculation_and_reporting_rules_OTV_raw_data.pdf)

##### **Publisher reach over several days**

In determining a publishers' reach over several days (e.g. week or month reach), SKO follows the rule that watching for the number of minutes that determine reach must occur on a single day.

For example: To be included in the 15 minute week reach a viewer must have watched the station concerned for 15 minutes on a single day. Watching a station for 5 minutes a day on three different days is not sufficient.



### 5.1.2. OTHER SKO RECOMMENDATIONS

#### **Overlap / Duplication Live**

The percentage of people who watched at least a specified portion of a video or in a time slot (a minimum percentage or number of minutes/seconds) and the same portion of another video or time slot. The broadcasts/time intervals slots do not have to follow one another.

For example: The specified minimum portion is 50%. An overlap of 45% for episode 1 of a new series with episode 2 means 45% that of those people who watched at least half of the first episode also watched at least 50% of the second episode.

#### **Selectivity index**

The relationship between an outcome in one target group with the outcome in the reference target group, multiplied by 100. This can be based on audience rating as well as on reach. The selectivity index for the reference target group is 100.

For example: A program produces an audience rating of 10% in the target group TOT 6+, 15% in the target group women 20-34. The selectivity index for the target group women 20-34 in relation to the target group TOT 6+ is 150.

#### **Population size**

SKO advises against using population sizes other than those published by SKO for the 53 standard target groups. The sample is not weighed for characteristics other than those that form the standard groups. If population sizes are to be estimated based on the audience research, then the daily sample for the first day that changes in population sizes or demographics become effective should be used for the estimation.



APPENDIX 1

# DEFINITIONS



## APPENDIX 1 DEFINITIONS

### 1. AUDIENCE RESEARCH FILES

#### 1.1 RESPONDENT DATA FILES

Data type	File Names	Example
Demographics	AUvX [YYMMDD]_calibrated.DEM	AUv4150621_calibrated.DEM
Devices	AUvX [YYMMDD]_calibrated.SET	AUv4150621_calibrated.SET
Video viewing	AUvX [YYMMDD]_calibrated.SWO	AUv4150621_calibrated.SWO
Video commercials	AUvX [YYMMDD]_vast_calibrated.CSV	AUv4150621_vast_calibrated.CSV

X = VERSION NUMBER OF THE FILE FORMAT (CURRENTLY V5)

.DEM = DEMOGRAPHICS

.SET = DEVICE/BROWSER INFORMATION

.SWO = VIEWING STATEMENTS

##### 1.1.1 VIEWING STATEMENTS

Online Video Usage, including when and for how long they watched, which device/browser they used and which station they watched. This information is available in viewing records. There are two files containing viewing records, video viewing (AUvXYMMDD\_calibrated.SWO) and online commercial viewing (AUvXYMMDD\_vast\_calibrated.CSV). They need to be combined with the files containing the MDU used equipment (Devices/Browsers information) (AUvXYMMDD\_calibrated.SET) and the demographic information through the Extended Household ID. Each viewing record represents an unaltered situation: the same person(s) watching the same station/publisher on the same device. When the situation changes (for example, when someone stops watching, someone new starts watching or the channel or publisher is changed), a new viewing record beings.

##### 1.1.2 BACKGROUND CHARACTERISTICS

The panellists in the calibrated AUvXYMMDD\_calibrated.DEM files represent an extended individual from the Media Panel. The extended household ID in position 165-174 of the file should be used to linked the different files.

The file AUvX.YYMMDD\_calibrated.DEM contains panel members' background characteristics as measured in the base and additional interviews or as establish at other moment during the year, which can be used to define target groups.

If the demographics file does contain '0' that are not documented in the documents SKO\_video\_format\_VX, that means that data is not available. This can be the case when interviews are yet to be taken to new panellists.

The demographic (as well as the viewing statements and MDU) files are issued daily at 13:00 hours. For the online commercials, the delivery of the demo files is at 10:00 hours and it is based on



overnights +1 (files over the weekend are delivered on Tuesday) and for the Online video content the delivery is based on overnights +10.

The demographic files delivered for video commercials are the same as for the online video programmes, they are delivered earlier (based on overnights +1 schedule).

**DELIVERY SCHEDULE - CALIBRATED COMMERCIALS**

Broadcast day			Delivery day	
Monday	8th	→	Wednesday	10th
Tuesday	9th	→	Thursday	11th
Wednesday	10th	→	Friday	12th
Thursday	11th	→	Saturday	13th
<b>Friday</b>	<b>12th</b>	→	<b>Tuesday</b>	<b>16th</b>
<b>Saturday</b>	<b>13th</b>	→	<b>Tuesday</b>	<b>16th</b>
<b>Sunday</b>	<b>14th</b>	→	<b>Tuesday</b>	<b>16th</b>

**DELIVERY SCHEDULE - CALIBRATED PROGRAMS**

Broadcast day			Delivery day	
Monday	8th	→	Friday	19th
Tuesday	9th	→	Saturday	20th
Wednesday	10th	→	Sunday	21th
Thursday	11th	→	Monday	22th
Friday	12th	→	Tuesday	23th
Saturday	13th	→	Wednesday	24th
Sunday	14th	→	Thursday	25th

1.1.3 WEIGHT FACTORS

The demographics file (AUvXYMMDD\_calibrated.DEM) contains the daily weight factor for the panel members. The weight factor is composed of 8 positions, of which the last 4 positions should be read as a decimal value (see SKO\_video\_format\_VX,). A projection factor, by which the number of panel members is projected on to population sizes, is already included in the weight factor. Thus, the sum of the weight factors for one day gives the population size. Panellists not having viewing on a particular day, are not present on the demographic files, hence, they do not have weights.

A number of individuals in the panel correspond to the non-internet population. These individuals all receive the viewer type =2 on the position 19 of the demographic files. These individuals have no viewing statements, but they should be accounted of when the weightings are applied, the sum of weightings is calculated.

The composition of the audience panel can vary daily because of panel members leaving and new members being added, technical problems, vacations, etc. To compensate for these fluctuations,



weighting must be carried out on a daily basis. As a result, each day, each respondent aged six and older has a separate weight factor.

### **IMPORTANT notice on weighting**

When computing ratings of the type online linear viewing (OLV), VOSDAL, online day of broadcast (OBD) as well as for online viewing moment report (OVM) and when calculating online commercials ratings, **the weights of the day of viewing** should be used.

When computing time shifted viewing ratings for online programmes, results types of Online Broadcast Total (OBT), Online Broadcast Total 28 days (OB28) or results from buckets 3 and 4, the weight of the panellists **on the BROADCAST day** should be used, instead of the weight on the day of viewing.

## 1.2. BROADCAST DATA FILES

### 1.2.1 PROGRAMS

The file `jjjjmmdd.PRL` contains information on all program types (these include programs, promotional videos, billboards, Postbus 51, advertising block, home shopping, station id's and static) broadcast on the coded stations (an up-to-date overview of all stations is available at [www.kijkonderzoek.nl](http://www.kijkonderzoek.nl)) on the day in question. In addition to the title, the starting and ending minute of the broadcast element and program type, this file contains the SKO codes. In addition, a variable is available that indicates which program segments belong together. There is also a serial number per broadcast element per station per day.

From 1-1-2011 on this file also includes an additional field for programme title. This field contains the programme title that the stations provide on their pre-log files. This title is not harmonized by The Nielsen Company. From 1-1-2011 more information on promos is available. New fields Promo Type Id, Promoted day, Promoted channel ID have been added. Programme titles are added in the files Secondary and Tertiary programme titles in case the promo refers to more than one channel. From 1-1-2012 the duration (in seconds) of promo's and other programme types on the PRL file has been added.

The link between programmes in the Online viewing files `AUvXYMMDD_calibrated.SWO` can be done by the PRLID. In the online programme viewing files ONLY the PRL-ID of the first segment of the programme on the day of broadcast is delivered in position 230-239. This PRL\_ID can be used to match with the Nielsen PRL ID of the broadcast day.

By matching to PRL file via the PRL-ID of the first segment it is possible to obtain the harmonized programme information (title, genre, etc..).

For the calculation of ratings, the software needs to use the **net duration of the programmes**. This can be obtained from in the Nielsen PRL file or in the calibrated online video data, in the so-called *online video clip length*. The net online video clip length is to be found as Duration in the position 106-



110 of the AUvXYMMDD\_calibrated.SWO file. This is the duration in seconds of the video (sko\_cl or clip length). In the Nielsen PRL,

Please note that the length of the viewing statement can be higher than the net length of the programme in the Nielsen file, or the length of the PRL\_ID of the first segment. This is due to re-viewing or the time elapsed due to pause .

A way to correct for this is to assign the duration of viewing (could be higher than program length) to different first segments of this broadcast. By that the first segment will receive a synthetic statements. The total duration of the clip (or the net total program duration in the PRL file) will be used to calculate the rating.

In the AUvXYMMDD\_calibrated.SWO file v5 also the duration in seconds of the first segment of the PRL is provided (in position X-X)

#### 1.2.2 CLASSIFICATION OF BROADCAST COMMERCIALS

The file jjjjmdd.cla contains information on the commercial broadcast on the day in question that can be used to classify the broadcast spots. Per TVTID, the file includes the link of the reference commercial, its harmonized title and harmonized information on the product advertised in the commercial (brand, sub-brand, product name, branch and advertiser).

Since 1-1-2011 and for billboards, this file can also contain multiple advertisers and products per TVTIDs. For this reason a new field has been added to the file (MultipleProdcut). Since 1-1-2011, the field ProductSequenceNumber has been deleted from the CLA file.

TVTID's have a product (combination of brand-subbrand-productname) and an advertiser assigned to it, which is presented in the CLA. Since the product's advertiser can change over time, due to companies take overs, the structure of the CLA-file changed from 1-7-2012. In the new file structure it is possible to assign a new advertiser to an existing product with a valid from date.

From 1-7-2012 the full list of TVTID's and their respective information, will be supplied in one file (the COMPLETECLA) on a monthly basis.

#### 1.2. 3. ONLINE COMMERCIALS

The file jjjjmdd.COL contains information on the online commercial broadcasts up to the broadcast day mentioned in the filename. Per ONLINEID (unique creative), the file includes the broadcast data of the Online commercial, its harmonized media agency and harmonized information on the product advertised in the online commercial (brand, sub-brand, product name, branch and advertiser).



This file contains always the whole history from 1-3-2015.

The Online commercials receive the same classification as the same creatives broadcasted on TV.

Nielsen is also providing a way to match the TV and Online commercials. This is done by a comparison of the commercial based on duration, audio and video. If the Creative's are exactly the same Nielsen adds the corresponding TVTID to the .COL file. A direct match between TV and Online creatives can be made by using the field nr 28 'TVTID'

The Online ID of the jjjjmmdd.COL can be linked to the calibrated online commercial viewing files by using the Online ID. This is provided in position 22 of the AUvXYMMDD\_vast\_calibrated.CSV file.

#### 1.2.4. ONLINE ONLY VIDEO CONTENT

Metadata on online only content is to be found in the panel data statements. From that data the following information can be derived:

- PROGRAMME NAME: Name or title of the video content or clip
- STREAM TYPE: Type of content (Linear, Episode or a segment of a programme)
- PUBLISHER
- GENRE
- ONLINE DATE: Date of actual online release of the clip

## 2. CALCULATION OF THE LENGTH OF THE VIEWING STATEMENT OF ONLINE PROGRAMMES FILES

The length of the viewing statement can be calculated from the START and END time provided in the AUvXYMMDD\_calibrated.SWO file. In order to have a standard for the calculation, we agree to subtract one second of all calculations. Viewing statements lengths can be calculated as:

$$=((\text{END HOUR} * 3600 + \text{END MIN} * 60 + \text{END SEC}) - (\text{START HOUR} * 3600 + \text{START MIN} * 60 + \text{START SECOND})) / 3600$$

OR

The difference between the ENDFHOUR-START HOUR, the difference between END MINUTE- START MINUE, the difference between the END SECOND - START SECOND and the formulat

$$=\text{HOUR DIFF} * 3600 + \text{MIN DIFF} * 60 + \text{SECOND DIFF}$$

This can be defined /summarized as extracting one second from the End second statement. For instance, if a statement END TIME is 20:35:15 AUvXYMMDD\_calibrated.SWO file, then the 15th second is not included in the length of the statement.





APPENDIX 2

# PERIOD SAMPLE CRITERIA



## APPENDIX 2 PERIOD SAMPLE CRITERIA

These criteria for period samples are used for the calculations of all relevant result types.

**Weeks:** The number of weeks in the period sample.

**Week days:** The maximum number of a specific day of the week that can be missing

**Total:** The maximum number of days in total that can be missing

TABEL 7.1

Weeks	Week-days	Total
1	0	0
2	1	1
3	1	1
4	1	2
5	1	2
6	1	2
7	1	2
8	1	3
9	1	3
10	1	3
11	1	3
12	1	3
13	1	4
14	1	4
15	1	4
16	1	5
17	1	5
18	1	5
19	1	6
20	1	6
21	2	6
22	2	6
23	2	6
24	2	7
25	2	7
26	2	8
27	2	8

Weeks	Week-days	Total
28	2	8
29	2	9
30	2	9
31	2	9
32	2	10
33	2	10
34	2	10
35	2	11
36	3	11
37	3	11
38	3	12
39	3	12
40	3	12
41	3	12
42	3	13
43	3	13
44	3	13
45	3	14
46	3	14
47	3	14
48	3	15
49	3	15
50	3	15
51	4	15
52	4	16
53	4	16
54	4	16



APPENDIX 3

**OVERVIEW OF VARIABLES  
NOT FOR EXTERNAL  
REPORTING**



### APPENDIX 3 OVERVIEW OF VARIABLES NOT FOR EXTERNAL REPORTING

A number of variables are only to be reported internally and for Q&C use. The following variables should not be used to external reporting:

- **SET data** = Browser, Coverage, Freq Usage Browser, Browser type, Browser ID, Operating System, Ownership, Main user and Main location
- **Demographics data** = Sector, Viewer Type, Sample type.
- **Viewing data** = Player, Activity type, Stream Type, is APP, Player version, MDI, MMI
- **Commercials data** = Creative ID, Browser ID, Client, Count, is App.

